



PCT09

ENTERED

RAW SEQUENCE LISTING

DATE: 09/18/2002

PATENT APPLICATION: US/09/980,372B

TIME: 16:51:20

Input Set : A:\PTO.DC.txt

Output Set: N:\CRF4\09182002\I980372B.raw

3 <110> APPLICANT: Aventis Research & Technologies GmbH & Co KG
 5 <120> TITLE OF INVENTION: Spliceosomal protein and its use
 7 <130> FILE REFERENCE: 199at09
 C--> 9 <140> CURRENT APPLICATION NUMBER: US/09/980,372B
 10 <141> CURRENT FILING DATE: 2000-05-03
 12 <150> PRIOR APPLICATION NUMBER: DE 19925668.3
 13 <151> PRIOR FILING DATE: 1999-06-04
 15 <160> NUMBER OF SEQ ID NOS: 18
 17 <170> SOFTWARE: PatentIn Ver. 2.1
 19 <210> SEQ ID NO: 1
 20 <211> LENGTH: 18
 21 <212> TYPE: PRT
 22 <213> ORGANISM: Homo sapiens
 24 <400> SEQUENCE: 1
 25 Lys Met Asn Ala Arg Thr Tyr Met Asp Val Met Arg Glu Gln His Leu
 26 1 5 10 15
 28 Thr Lys
 32 <210> SEQ ID NO: 2
 33 <211> LENGTH: 15
 34 <212> TYPE: PRT
 35 <213> ORGANISM: Homo sapiens
 37 <400> SEQUENCE: 2
 38 Lys Leu Thr Ala Thr Pro Thr Pro Leu Gly Gly Met Thr Gly Phe
 39 1 5 10 15
 42 <210> SEQ ID NO: 3
 43 <211> LENGTH: 10
 44 <212> TYPE: PRT
 45 <213> ORGANISM: Homo sapiens
 47 <400> SEQUENCE: 3
 48 Lys Ala Ile Val Asn Val Ile Gly Met His
 49 1 5 10
 52 <210> SEQ ID NO: 4
 53 <211> LENGTH: 8
 54 <212> TYPE: PRT
 55 <213> ORGANISM: Homo sapiens
 58 <400> SEQUENCE: 4
 59 Lys Arg Ile Phe Glu Ala Phe Lys
 60 1 5
 63 <210> SEQ ID NO: 5
 64 <211> LENGTH: 12
 65 <212> TYPE: PRT
 66 <213> ORGANISM: Homo sapiens
 68 <400> SEQUENCE: 5

PATENT APPLICATION: US/09/980,372B

TIME: 16:51:20

Output Set: N:\CRF4\09182002\I980372B.raw

```

69 Lys Leu Arg Arg Met Asn Arg Phe Thr Val Ala Glu
70 1 5 10
73 <210> SEQ ID NO: 6
74 <211> LENGTH: 15
75 <212> TYPE: PRT
76 <213> ORGANISM: Homo sapiens
78 <400> SEQUENCE: 6
79 Lys Arg Thr Gly Ile Gln Glu Met Arg Glu Ala Leu Gln Glu Lys
80 1 5 10 15
83 <210> SEQ ID NO: 7
84 <211> LENGTH: 12
85 <212> TYPE: PRT
86 <213> ORGANISM: Homo sapiens
88 <400> SEQUENCE: 7
89 Lys Leu Thr Ile His Gly Asp Leu Tyr Tyr Glu Gly
90 1 5 10
93 <210> SEQ ID NO: 8
94 <211> LENGTH: 16
95 <212> TYPE: PRT
96 <213> ORGANISM: Homo sapiens
98 <400> SEQUENCE: 8
99 Lys Leu Gly Ala Val Phe Asn Gln Val Ala Phe Pro Leu Gln Tyr Thr
100 1 5 10 15
103 <210> SEQ ID NO: 9
104 <211> LENGTH: 10
105 <212> TYPE: PRT
106 <213> ORGANISM: Homo sapiens
108 <400> SEQUENCE: 9
109 Lys Leu Leu Arg Val Tyr Asp Leu Gly Lys
110 1 5 10
113 <210> SEQ ID NO: 10
114 <211> LENGTH: 16
115 <212> TYPE: PRT
116 <213> ORGANISM: Homo sapiens
118 <400> SEQUENCE: 10
119 Lys Asn Val Ser Glu Glu Leu Asp Arg Thr Pro Pro Glu Val Ser Lys
120 1 5 10 15
123 <210> SEQ ID NO: 11
124 <211> LENGTH: 11
125 <212> TYPE: PRT
126 <213> ORGANISM: Homo sapiens
128 <400> SEQUENCE: 11
129 Lys Leu Glu Asn Ile Ala Gln Arg Tyr Ala Phe
130 1 5 10
133 <210> SEQ ID NO: 12
134 <211> LENGTH: 13
135 <212> TYPE: PRT
136 <213> ORGANISM: Homo sapiens
138 <220> FEATURE:

```

RAW SEQUENCE LISTING

DATE: 09/18/2002

PATENT APPLICATION: US/09/980,372B

TIME: 16:51:20

Input Set : A:\PTO.DC.txt

Output Set: N:\CRF4\09182002\I980372B.raw

```

139 <221> NAME/KEY: UNSURE
140 <222> LOCATION: (8)
142 <400> SEQUENCE: 12
W--> 143 Lys Val Ser Glu Pro Leu Leu Xaa Glu Leu Phe Leu Gln
144   1           5           10
147 <210> SEQ ID NO: 13
148 <211> LENGTH: 20
149 <212> TYPE: PRT
150 <213> ORGANISM: Homo sapiens
152 <400> SEQUENCE: 13
153 Lys Asp Arg Val Thr Gly Gln His Gln Gly Tyr Gly Phe Val Glu Phe
154   1           5           10           15
156 Leu Ser Glu Glu
157           20
160 <210> SEQ ID NO: 14
161 <211> LENGTH: 7
162 <212> TYPE: PRT
163 <213> ORGANISM: Homo sapiens
165 <400> SEQUENCE: 14
166 Lys Glu Tyr Asp Pro Leu Lys
167   1           5
170 <210> SEQ ID NO: 15
171 <211> LENGTH: 11
172 <212> TYPE: PRT
173 <213> ORGANISM: Homo sapiens
175 <400> SEQUENCE: 15
176 Lys Arg Trp Arg Thr Arg Val Trp Asp Asn Asp
177   1           5           10
180 <210> SEQ ID NO: 16
181 <211> LENGTH: 22
182 <212> TYPE: RNA
183 <213> ORGANISM: Homo sapiens
185 <400> SEQUENCE: 16
186 auuuuccuua cucauaagdd dd 22
189 <210> SEQ ID NO: 17
190 <211> LENGTH: 1067
191 <212> TYPE: DNA
192 <213> ORGANISM: Homo sapiens
194 <220> FEATURE:
195 <221> NAME/KEY: CDS
196 <222> LOCATION: (214)..(954)
198 <400> SEQUENCE: 17
199 ctgacatcag gagtttgagg ccggcttgga acatgggtgaa atcctgtctg tactagaaat 60
201 gcaaaaatta gctgggcgtg gtgggtgtgtg tctgtgatcc cagctgctcg gcctcccaag 120
203 gtgctgggat tacaggcgtg agccaccgcg tctggcctca gccaaaggttt ttaagtaaca 180
205 tatttcagca ttggctctac agcgttgacag aac atg aac gat tgg atg ccc atc 234
206                               Met Asn Asp Trp Met Pro Ile
207                               1           5
209 gcc aag gag tat gat cca ctc aaa gcg ggc agc att gat ggc acc gat 282

```

RAW SEQUENCE LISTING

DATE: 09/18/2002

PATENT APPLICATION: US/09/980,372B

TIME: 16:51:20

Input Set : A:\PTO.DC.txt

Output Set: N:\CRF4\09182002\I980372B.raw

```

210 Ala Lys Glu Tyr Asp Pro Leu Lys Ala Gly Ser Ile Asp Gly Thr Asp
211      10      15      20
213 gaa gac cca cac gac cgc gcg gtc tgg agg gca atg ctg gca cga tat 330
214 Glu Asp Pro His Asp Arg Ala Val Trp Arg Ala Met Leu Ala Arg Tyr
215      25      30      35
217 gtc ccc aac aaa ggt gtc ata gga gat ccc ctc ctc acc ctg ttt gtg 378
218 Val Pro Asn Lys Gly Val Ile Gly Asp Pro Leu Leu Thr Leu Phe Val
219      40      45      50      55
221 gcc aga cta aac ttg cag acc aag gag gac aaa tta aag gaa gtc ttt 426
222 Ala Arg Leu Asn Leu Gln Thr Lys Glu Asp Lys Leu Lys Glu Val Phe
223      60      65      70
225 tcc cgc tat ggt gac atc cgg cgg ctt cgg ctg gtc agg gac ttg gtc 474
226 Ser Arg Tyr Gly Asp Ile Arg Arg Leu Arg Leu Val Arg Asp Leu Val
227      75      80      85
229 aca ggt ttt tca aag ggc tac gcc ttc atc gaa tac aag gag gag cgt 522
230 Thr Gly Phe Ser Lys Gly Tyr Ala Phe Ile Glu Tyr Lys Glu Glu Arg
231      90      95      100
233 gcc gtg atc aaa gct tac cga gat gct gat ggc ctg gtt att gac cag 570
234 Ala Val Ile Lys Ala Tyr Arg Asp Ala Asp Gly Leu Val Ile Asp Gln
235      105      110      115
237 cat gag ata ttt gtg gac tac gag ctg gaa agg act ctc aaa ggg tgg 618
238 His Glu Ile Phe Val Asp Tyr Glu Leu Glu Arg Thr Leu Lys Gly Trp
239      120      125      130      135
241 atc cct cgg cga ctt gga ggc ggt ctt ggg gga aaa aag gag tct ggg 666
242 Ile Pro Arg Arg Leu Gly Gly Gly Leu Gly Gly Lys Lys Glu Ser Gly
243      140      145      150
245 caa ctg aga ttt ggg gga cgg gac cgg cct ttt cga aaa cct att aac 714
246 Gln Leu Arg Phe Gly Gly Arg Asp Arg Pro Phe Arg Lys Pro Ile Asn
247      155      160      165
249 ttg cca gtt gtt aaa aac gac ctc tat aga gag gga aaa cgg gaa agg 762
250 Leu Pro Val Val Lys Asn Asp Leu Tyr Arg Glu Gly Lys Arg Glu Arg
251      170      175      180
253 cgg gag cga tct cga tcc cga gaa aga cac tgg gac tcg agg aca agg 810
254 Arg Glu Arg Ser Arg Ser Arg Glu Arg His Trp Asp Ser Arg Thr Arg
255      185      190      195
257 gat cga gac cat gac agg ggc cgg gag aag aga tgg caa gaa aga gag 858
258 Asp Arg Asp His Asp Arg Gly Arg Glu Lys Arg Trp Gln Glu Arg Glu
259      200      205      210      215
261 ccg acc agg gtg tgg ccc gac aat gac tgg gag aga gag agg gac ttc 906
262 Pro Thr Arg Val Trp Pro Asp Asn Asp Trp Glu Arg Glu Arg Asp Phe
263      220      225      230
265 aga gat gac agg atc aag ggg agg gag aag aag gaa aga ggc aag tag 954
266 Arg Asp Asp Arg Ile Lys Gly Arg Glu Lys Lys Glu Arg Gly Lys
267      235      240      245
269 aggcccaaca gcagaacccc aaagtgaagt tacagtggaa atgagtggag ggggattgtc 1014
271 tttcaacgca gcgtgagtct aatggttgaa taaaacttac tgatgatcaa aaa 1067
274 <210> SEQ ID NO: 18
275 <211> LENGTH: 246
276 <212> TYPE: PRT

```

RAW SEQUENCE LISTING

DATE: 09/18/2002

PATENT APPLICATION: US/09/980,372B

TIME: 16:51:20

Input Set : A:\PTO.DC.txt

Output Set: N:\CRF4\09182002\I980372B.raw

```

277 <213> ORGANISM: Homo sapiens
279 <400> SEQUENCE: 18
280 Met Asn Asp Trp Met Pro Ile Ala Lys Glu Tyr Asp Pro Leu Lys Ala
281   1           5           10           15
282 Gly Ser Ile Asp Gly Thr Asp Glu Asp Pro His Asp Arg Ala Val Trp
283           20           25           30
284 Arg Ala Met Leu Ala Arg Tyr Val Pro Asn Lys Gly Val Ile Gly Asp
285           35           40           45
286 Pro Leu Leu Thr Leu Phe Val Ala Arg Leu Asn Leu Gln Thr Lys Glu
287   50           55           60
288 Asp Lys Leu Lys Glu Val Phe Ser Arg Tyr Gly Asp Ile Arg Arg Leu
289   65           70           75           80
290 Arg Leu Val Arg Asp Leu Val Thr Gly Phe Ser Lys Gly Tyr Ala Phe
291           85           90           95
292 Ile Glu Tyr Lys Glu Glu Arg Ala Val Ile Lys Ala Tyr Arg Asp Ala
293           100          105          110
294 Asp Gly Leu Val Ile Asp Gln His Glu Ile Phe Val Asp Tyr Glu Leu
295           115          120          125
296 Glu Arg Thr Leu Lys Gly Trp Ile Pro Arg Arg Leu Gly Gly Gly Leu
297   130          135          140
298 Gly Gly Lys Lys Glu Ser Gly Gln Leu Arg Phe Gly Gly Arg Asp Arg
299  145          150          155          160
300 Pro Phe Arg Lys Pro Ile Asn Leu Pro Val Val Lys Asn Asp Leu Tyr
301           165          170          175
302 Arg Glu Gly Lys Arg Glu Arg Arg Glu Arg Ser Arg Ser Arg Glu Arg
303           180          185          190
304 His Trp Asp Ser Arg Thr Arg Asp Arg Asp His Asp Arg Gly Arg Glu
305           195          200          205
306 Lys Arg Trp Gln Glu Arg Glu Pro Thr Arg Val Trp Pro Asp Asn Asp
307   210          215          220
308 Trp Glu Arg Glu Arg Asp Phe Arg Asp Asp Arg Ile Lys Gly Arg Glu
309  225          230          235          240
310 Lys Lys Glu Arg Gly Lys
311           245

```

RAW SEQUENCE LISTING ERROR SUMMARY DATE: 09/18/2002
PATENT APPLICATION: US/09/980,372B TIME: 16:51:21

Input Set : A:\PTO.DC.txt
Output Set: N:\CRF4\09182002\I980372B.raw

Please Note:

Use of n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> to <223> fields of each sequence which presents at least one n or Xaa.

Seq#:12; Xaa Pos. 8

VERIFICATION SUMMARY

PATENT APPLICATION: US/09/980,372B

DATE: 09/18/2002

TIME: 16:51:21

Input Set : A:\PTO.DC.txt

Output Set: N:\CRF4\09182002\I980372B.raw

L:9 M:270 C: Current Application Number differs, Replaced Current Application Number

L:143 M:258 W: Mandatory Feature missing, <223> not found for SEQ ID#:12

L:143 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:12 after pos.:0